

Model Structure Validation of Cell Signaling Pathways Using Colored Petri nets

Dvid Csercsik* Katalin M. Hangos*

* *Process Control Research Group, Systems and Control Laboratory,
Computer and Automation Research Institute,
Hungarian Academy of Sciences
P.O. Box 63, H-1518 Budapest, Hungary
e-mail: {csercsik, hangos}@scl.sztaki.hu*

Abstract: In this paper we propose a novel colored timed Petri net modelling approach for model structure validation of cell signaling pathways. The aim is to distinguish between models of different structures based on typically qualitative, roughly detailed measurement data. In this approach, tokens are present all the time at all of the places, and their color corresponds to the actual quantized activation state of the signaling component. The activation states are updated according to the progression of the signal through the signaling network. We demonstrate the results on a signaling network, which describes the interaction of fast (G protein coupled) and slow (β -arrestin coupled) transmission.

Keywords: Modeling and Identification; Control Applications in Biological Processes

1. INTRODUCTION

Implied by the recently revealed extreme amount of biological data, the need to describe the structure and functionality of intracellular biochemical networks provided a wide apparatus of applied mathematical and computational methods [Fisher and Henzinger, 2007, Gilbert et al., 2006]. The ordinary differential equation (ODE) approach, which is probably the most popular, together with reaction-diffusion type PDE-s, gave birth to valuable predictions and explanations in many different topics [Bhalla et al., 2002, Craciun and Feinberg, 2006, Huang and Ferrell Jr., 1996, Ferrel Jr., 1996, Tyson et al., 2003].

Besides of the ODE based dynamical descriptions, Petri nets [Angeli et al., 2007, Oliveira, 2001] form a class of computational models, which are strongly based on the interactions in a straightforward way, but also include a state space (dominantly finite), which makes them able to describe the time evolution of signaling events. If this state space is defined with an appropriate resolution, it can be easily bound to qualitative interpretation of the corresponding measurement data. In this paper we propose a Petri net modelling approach for the description of intracellular signaling networks that is used for model structure validation.

As discussed in [Heiner et al., 2004], typically in the biological literature, the various descriptions of a certain signaling pathway differ not only in the abstraction level, but also in their focus. Therefore, the modelling procedure may lead to multiple feasible models with different struc-

ture. To distinguish between models of various structures, model validation is needed, which on one hand can be based on the analysis of model properties (see eg. [Heiner et al., 2004]), and on the other hand on comparison with measurement data if available. In this paper we will focus on the second way of model validation to solve the inverse problem of modelling and we suppose that some kind of experimental measurement data of sparse granularity are available for model validation, which can be interpreted by the output of our model.

2. COLOURED PETRI NET DESCRIPTION OF SIGNALING PATHWAYS

2.1 Coloured Petri nets

Petri nets (PNs) [Murata, 1989] were introduced as graphical tools for the description and analysis of concurrent processes which arise in systems with many components. The structure of Petri nets is a bipartite graph, its nodes can only be places (denoting substances, states or conditions) and transitions (denoting reactions, state transformations) connected into a bipartite directed graph. Places are usually depicted as circles, while transitions as full rectangles in the figures.

Formally speaking, a coloured Petri net [Jensen, 1997] is bipartite multigraph described by a 6-tuple

$$\langle P, T, C, I^-, I^+, M_0 \rangle$$

where P is a set of places, T is a set of transitions and P and T are disjoint. C is a color function defined from P into a finite and non-empty sets. I^- and I^+ are the backward and forward incidence functions defined on $(P \times T)$ that $I^-(p, t), I^+(p, t) \in [t \mapsto C(p)], \forall (p, t) \in (P \times T)$. A marking

* This work was supported in part by the Hungarian Research Fund through grant 83440, and by the by the Control Engineering Research Group of the Budapest University of Technology and Economics

is a distribution of coloured tokens on the places of the CPN described by a marking function M defined on P . For each place p_i a finite set of colors C_{p_i} is given that determine the color of the tokens appearing on that place. An initial marking $M_0(p) \in C(p), \forall p \in P$ is given for each CPN. Arcs representing I^- and I^+ are often called input arcs and output arcs, respectively. These have a weight, called transcription describing the actual function they represent. Transitions may also have inscriptions that determine their operation.

A transition is enabled (it may fire) in M if there are enough tokens in its input places (noted by p) for the consumptions to be possible, i.e. iff $\forall p: M(p) > I^-(p, t)$, and if their distribution corresponds to a row in the *activation table*, an inscription of the transition. The input places of a transition are called conditions, and the output places are called consequences. A firing time may also be associated to each transition in timed Petri nets. Note that enabled transitions may fire in arbitrary order. The operation of a CPN is as follows: firing a transition t in a marking M consumes $I^-(p, t)$ tokens from each of its input places $p \in P$, and produces $I^+(p', t)$ tokens in each of its output places $p' \in P$ the color of which is determined by the *activation function*.

2.2 Signaling networks and Petri nets

One widespread way of describing metabolic and signalling networks by Petri nets is based on the assumption, that one token represents one molecule or one chemical group in the reaction network, thus this approach is able to give a realistic description of the system on the molecular level. In this approach, on which vast majority of literature corresponding to Petri net modelling of biochemical pathways is based, the Petri net representation of an interaction is based on its molecular mechanism [Oliveira, 2001, Reddy et al., 1993].

In contrast to metabolic networks, the elements of signaling networks are constantly present in the system (at each stage of the signaling process), and their main feature that is influenced by the various input signals, is the activation state. From the point of view of molecular biology, the activation state can also be physically represented - in most cases by the by phosphorylated-dephosphorylated state of the molecule -, so the above mentioned molecular mechanism based description may be used also in the case of signaling networks [Li et al., 2006, 2007, Hardy, 2004, Grunwald et al., 2008], including the case of hybrid Petri nets [Matsuno et al., 2003]. In this approach the number of tokens may correspond to activation level of certain components [Li et al., 2007].

At the same time, the number of places, transitions and simulation steps may be very high in the case of complex signaling networks. Furthermore, a situation of combined activation and inhibition can appear in cell signaling pathways, the description of which is not straightforward with the molecular level approach, may lead to an oscillating behavior of the model.

Colored Petri Nets (CPNs) have also been applied for the simulation of signaling pathways. In [Lee et al., 2006] CPNs are used as a tool for the approxima-

tion/computational implementation of the underlying dynamic equations describing the time evolution of concentrations of components in the system. The approach detailed in this paper is based on the molecular mechanisms as well, which determines the kinetic equations via the mass action law, and these ODE's are implemented with colored Petri nets. Timed continuous Petri nets are also discussed in [Vazquez et al., 2011].

2.3 Coloured Petri net model of signaling systems

Biochemical signaling pathways are fundamentally based on regulatory enzymatic phosphorylation-dephosphorylation type activation-inactivation protein interactions. Here we describe the quantized *qualitative activation state* of the particular elements by specific colors of a *finite color set* \mathcal{C}

$$\mathcal{C} = \{0, S, N, L\} \quad (1)$$

corresponding to zero (0), small (S), normal (N) and large (L) actual activity. We suppose that these colors refer to the ratio of the activated/nonactivated proteins, or the relative level of certain signaling elements (eg. cAMP).

This detail of activation levels makes *the model output comparable with qualitative measurement results (eg. western blot)*, where only similar detail of activation levels can be distinguished. The above approach means that we will not necessary use the detailed molecular mechanism-based description, but rather a different simplified approach.

Modelling assumptions Based on the assumption that the components of a signaling network are always present in the system, the most important novel feature of the proposed model is that always exactly one token is located on each place, and the tokens' colours are updated according to activation functions and activation states of upstream states (components whose activation influences the activation state of the actual component). Formally speaking, this means, that all places are both conditions and consequences of the same transition at the same time as in [Lee et al., 2006]. This implies that each transition is enabled and can fire in each discrete time interval.

Transitions are associated to the reactions of the signalling system. The *difference in the reaction rates are described in the firing time of the individual transitions*.

Activation functions The activation function defines the reaction of an effector and a target. This function will describe eg. that if component A influences (eg. activates) component B, then the state of A will not depend on the state of B during the transition, while the reverse is not true. An activation function is a discrete function that formally describes the mapping $T_{k+1} = T_k(A_k)$, where T_k is the qualitative value of the activation state of the target in the time instance k , and A_k is that of the activator. As both $T - k$ and A_k are discrete values from \mathcal{C} in Eq. (1), it can be given in the form of a lookup table.

Using the above activation functions, an activation reaction (eg. protein phosphorylation) can be described by a transition depicted in Fig. 1 in the colored Petri net. If we

assume that the activation state of the activator remains unchanged during the simulation, the activation state of the target will evolve in time as depicted in Fig. 2.

It is important to note that *the activation function in Fig. 1 is not unique*. Any discrete function $\text{Fun_act}(A, T)$ corresponds to an activation reaction, when

- (i) the activation state of the target increases if the activation state of the activator is greater than itself, i.e. $T_{k+1} > T_k$ when $A_k > T_k$, and
- (ii) the activation state of the activator remains unchanged, i.e. $A_{k+1} = A_k$

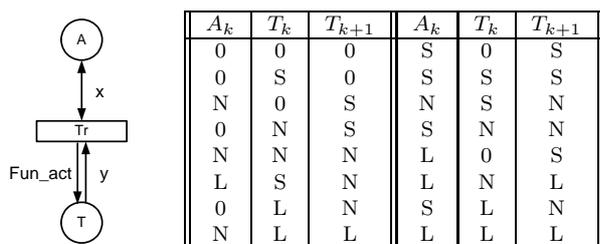


Fig. 1. Proposed Petri net representation of an activation (eg. phosphorylation) reaction, and a possible activation function, defined by a lookup table. The transition Tr checks the current activation state of the activator - eg. a kinase - (A) and the target - protein - (T) in every time step, and according to a function ($\text{Fun_act}(A_k, T_k)$), updates the activation state of the target, if necessary. The activation state of the activator is not influenced.

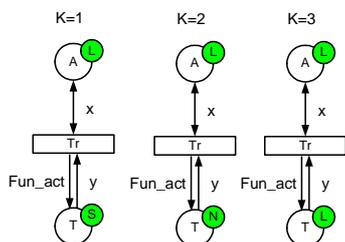


Fig. 2. Proposed Petri net representation of an activation (eg. phosphorylation) reaction. The green circles represent the tokens. The states evolve according to $\text{Fun_act}(A, T)$ described in Table in Fig. 1

Convergence and divergence Since our approach assumes that tokens are present in each place at every time point, the concepts of convergence and divergence have to be reconsidered. If we consider one target element in the network which is influenced by two activators, or by an activator and an inhibitor (eg. a kinase and a phosphatase), the Petri net representation of **convergence** will be straightforward as depicted in Fig. 3. The *convergent point* of the signaling pathway model is described with functions of multiple arguments (eg. 3 argument in the case of 2 activator - 2 activator states and one state corresponding to the target). Via the corresponding activation function it can be defined eg. whether the activity of both activators is necessary for a certain level of the target, or one of them is enough.

In addition, if the activation function is determined in a proper way, oscillatory behavior of the model can be avoided in the case of parallel activation and inhibition,

and the state trajectory will converge to an equilibrium. One good example for this case can be the adenylate-cyclase (AC), the activity of which is enhanced by $G_s\alpha$ proteins, and inhibited by $G_i\alpha$ proteins [Purves et al., 1997]. Higher degree convergence points (with more than 2 activators/inhibitors) can be implemented similarly by functions with more arguments, or can be composed in a hierarchical way of convergence points with lower degree, and places representing instrumental variables.

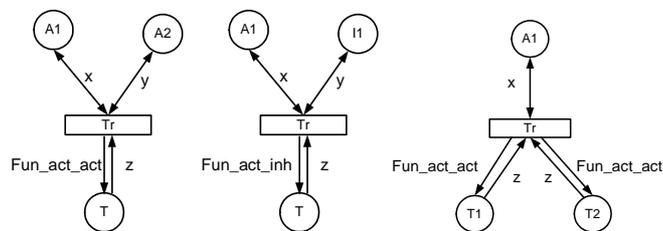


Fig. 3. Petri net representation of convergence and divergence in the proposed methodology. Convergence: combined activation-activation (eg. multisite phosphorylation by different enzymes) and activation-inhibition (eg. by a kinase and a phosphatase). Divergence: The number of output places is determined by the number of elements in the system which are influenced by the activator A1. The color of the multiple output places (consequences) will always be the same during simulation.

We talk about **divergence** in the signaling pathway, if the activation of an effector directly influences the activation state of more than one target in the system, while the effector is not affected by the targets. In this case we define more output places for the transition. Each output place will hold a token of the same color, and each will activate a different corresponding downstream pathway¹. This structure is depicted in Fig. 3.

Timing and firing of transitions The kinetic model parameters are given in terms of the firing time of the transitions and the properties of the activation functions. The firing time of a transition can be defined based on the data available about the kinetics and molecular properties of the reaction associated with that transition. Simple phosphorylation reactions are assumed to be fast, while synthesis of second messengers as cAMP or recruitment of scaffolding proteins in the case of a multi-component signaling complex are considered to take more time.

The sampling time, that is the chosen timing of the model should fit to the time resolution of the available measurements. We assume that a transition fires each time when the color of the output differs from the one implied by the actual condition states.

If the presumed possible model structures (derived from the pathway maps) and the time delays of interactions (based on the molecular mechanism and kinetic properties of the reactions) are given, the final question is whether we are able to find feasible activation functions which describe

¹ Although the expressions *upstream* and *downstream* are widespread in the biomedical literature when describing signaling pathways, their usage may be misleading in the case when circles appear in the signaling map. In this article we will use these expressions in the terms of local relations, as synonyms of *input/output* and *condition/consequence* terminology of Petri-nets.

the measurement data. Feasibility in this case means the limitations of the possible activation functions (eg. an activator can not reduce its target's activity, a target in general can not affect the activation state of the enzyme acting on him etc.). Taking into account the concept of no/basal/moderate/high activation of components, and the implied size of the state space, one can identify the activation functions from measured data. However, intuitive approaches can be used, as well.

3. FAST AND SLOW TRANSMISSION: A CASE STUDY

In this section we apply the proposed modelling framework for the description of the interplay of rapid (G protein coupled) and slow (β -arrestin coupled) transmission in the signaling process of ERK (Extracellular regulated kinase) activation. We compare the results obtained via the proposed colored Petri net modelling method with results based on ODE description of the process [Csercsik et al., 2008].

3.1 Rapid (G protein dependent) and slow (β -arrestin dependent) transmission

The interplay of the two pathways on which ERK activation may take place has been described in [DeWire et al., 2007]. ERK may be activated in a rapid, transient way through G protein dependent signaling, and in a more slow but sustained manner through the formation of a β -arrestin dependent signaling complex formed from the phosphorylated receptor and scaffolding proteins.

G protein dependent ERK activation G protein-coupled receptors (GPCRs) reside in the membrane of the target cells, and after binding with the agonist ligand, they subsequently interact with their respective G proteins to induce a cascade of downstream i.e. intracellular signaling.

The G proteins are composed of α , β and γ subunits, which dissociate upon receptor-induced exchange of GDP for GTP on the α subunit ($G\alpha$) to form a free $G\alpha$ and a dimer of $G\beta\gamma$ subunits [Gutkind, 1998]. These $G\alpha$ subunits, as well as the dissociated $\beta\gamma$ subunits, and other receptor-interacting proteins are capable of initiating diverse downstream signaling pathways via second messenger molecules, such as cyclic AMP, cyclic GMP, inositol triphosphate, diacylglycerol, and calcium, depending eg. on the cellular context. In this work we focus on the $G\alpha$ dependent activation of ERK, $G\beta\gamma$ signaling events are not examined.

β -arrestin dependent ERK activation Following GPCR activation, the ligand-bind receptor can be phosphorylated by GPCR kinases (GRKs). As described for eg. in [Beaulieu et al., 2005] in the case of dopamine receptors, β -arrestins bind to the receptors after phosphorylation to uncouple them from G proteins and participate in the recruitment of the endocytic protein complex, thus leading to an attenuation of GPCR signaling.

On the other hand, a signaling complex binding to the phosphorylated receptor composed of β -Arrestin, ERK1/2, Raf-1 and MEK can initiate ERK activation [DeWire et al., 2007].

Together, the G-protein and β -arrestin coupled pathways form a signaling network convergent to the central target kinase ERK.

Regulation of G protein signaling Another important mechanism contributing to the dynamics of signaling is the feedback regulation via the RGS proteins (regulators of G protein signaling) [Krauker et al., 2002].

3.2 The CPN models of (G protein dependent) and slow (β -arrestin dependent) transmission

For the description of the signaling events related to ERK activation, we developed 3 different timed CPN models, corresponding to G protein signaling, β -arrestin dependent slow transmission and the combination of the two pathways respectively. All models were implemented in CPN Tools (<http://www.daimi.au.dk/CPNtools/>).

The right downward path in Fig. 4 (a), depicting the first model, describes the binding of the ligand to the G protein coupled receptor, $G\alpha$ activation, and $G\alpha$ mediated ERK activation, while the left upward path describes the activation of RGS proteins (left upward path) which regulate G protein signaling.

Fig. 4 (b) figure corresponds to the second model (slow transmission) and describes the formation of the signaling complex from the activated, ligand bound receptor and β -arrestin (receptor phosphorylation is not explicitly included).

Fig. 5 depicts the third model corresponding to the combination of the two pathways. As it can be seen in the figure, the complex formation process affects the state of the free ligand-bound GPCR (as described in section 3.1.1, if a signaling complex is formed the corresponding receptor no longer takes place in $G\alpha$ activation). On the bottom of the 5, the coactivation of ERK can be seen by $G\alpha$ -GTP and the β -arrestin dependent signaling complex. This can be considered as a typical example of convergence in signaling.

The simulation results of the three models can be seen in Fig. 6. To make the discrete levels of the CPN model comparable to measurement results and ODE simulation data, we depicted the values 0,S,N,L as 0, 0.2, 0.7 and 0.95 respectively.

As we can see, the first CPN model describing the G-protein dependent ERK activation and RGS signaling describes the fast transient activation of ERK, while the second model describes the sustained component. The third model, corresponding to combined activation describes appropriately the qualitative features of the measured ERK activation pattern.

3.3 Discussion

We can see that the qualitative features of the trajectories are quite well captured with the third proposed CPN model describing the combined activation of the pathways. Compared to the ODE model described in [Csercsik et al., 2008], holding 22 state variables and 46 parameters, the model proposed in this paper inhibits only 9 places, which can be considered as state variables, and 7 functions which

has been demonstrated on the model of fast and slow transmission. The model is able to describe the (dominantly phosphorylation-dephosphorylation based) activation-deactivation reactions typical in signaling networks, furthermore the convergence and divergence and of signaling pathways, including the case of combined activation-inhibition.

The defined model class provides a valuable tool for experiment-based model validation in order to distinguish between various models of different structure, representing different interpretations of biological data available about the signaling system of interest.

The modeling approach is demonstrated in the case of $G\alpha$ -GTP and β - *arrestin* dependent signaling in the case of the detailed case study. The case study demonstrates the implementation of convergence (ERK coactivation) and feedback (RGS dependent $G\alpha$ regulation) in the proposed modeling framework.

Taking into account the detail and variation of the measurement data compared to simulation results, the simple CPN model gives a quite good approximation of the qualitative features of the previously published complex ODE model [Cserssik et al., 2008] of the signaling events.

The state space of the model is able to distinguish between zero, low, normal and high activation state, which makes the detail of the results comparable to western blot and proteomics data. Due to the detailed simplifications, the number of simulation steps and the needed computational effort regarding model simulations is low.

REFERENCES

- D. Angeli, P. De Leenheer, and E.D. Sontag. A Petri net approach to the study of persistence in chemical reaction networks. *Mathematical Biosciences*, 210:598–618, 2007.
- J. M. Beaulieu, T. D. Sotnikova, S. Marion, R. J. Lefkowitz, R. R. Gainetdinov, and M. G. Caron. An Akt/ β -Arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior. *Cell*, 122:261–273, 2005.
- U. S. Bhalla, P.T. Ram, and R. Iyengar. MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. *Science*, 297:1018–1023, 2002.
- G. Craciun and M. Feinberg. Understanding bistability in complex enzyme-driven reaction networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103:8697–8702, 2006.
- D. Cserssik, K.M. Hangos, and G.M. Nagy. A simple reaction kinetic model of rapid (G protein dependent) and slow (β -Arrestin dependent) transmission. *Journal of Theoretical biology*, 255:119–128, 2008.
- S. M. DeWire, S. Ahn, R. J. Lefkowitz, and S. K. Shenoy. β -Arrestins and cell signaling. *Annual Reviews of Physiology*, 69:483–510, 2007.
- J.E. Ferrel Jr. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *TIBS*, 21:460, 1996.
- J. Fisher and T.A. Henzinger. Executable cell biology. *Nature Biotechnology*, 25:1239–1249, 2007.
- D. Gilbert, H. Fuß, X. Gu, R. Orton, S. Robinson, V. Vyshemirsky, M.J. Kurth, C.S. Downes, and W. Dunitz. Computational methodologies for modelling, analysis and simulation of signalling networks. *Briefings in Bioinformatics*, 7:339–353, 2006.
- S. Grunwald, A. Speer, Jörg Ackermann, and I. Koch. Petri net modelling of gene regulation of the duchenne muscular dystrophy. *BioSystems*, 92:189–205, 2008.
- J.S. Gutkind. The pathways connecting G protein-coupled receptors to the nucleus through divergent mitogen-activated protein kinase cascades. *Journal of Biological Chemistry*, 273:1839–1842, 1998.
- S. Hardy. Modeling and simulation of molecular biology systems using petri nets: Modeling goals of various approaches. *Journal of Bioinformatics and Computational Biology*, 2:619–637, 2004.
- M. Heiner, Ina Koch, and Jrgen Will. Model validation of biological pathways using petri nets-demonstrated for apoptosis. *BioSystems*, 75:15–28, 2004.
- C. Y. F. Huang and J. E. Ferrell Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the USA*, 93:10078, 1996.
- Kurt Jensen. *Coloured Petri nets*, volume 1. Springer Berlin / Heidelberg, 1997.
- D.C. Krauker, K.M Page, and S. Sealfon. Module dynamics of the GnRH signal transduction network. *Journal of Theoretical Biology*, 218:457–470, 2002.
- D.Y. Lee, R. Zimmer, S.Y. Lee, and S. Park. Colored Petri net modeling and simulation of signal transduction pathways. *Metabolic Engineering*, 8:112–122, 2006.
- C. Li, S. Suzuki, Q.W. Ge, and M. Nakata. Structural modeling and analysis of signaling pathways based on Petri nets. *Journal of Bioinformatics and Computational Biology*, 4:1119–1140, 2006.
- C. Li, Q.W. Ge, M. Nakata, H. Matsuno, and S. Miyano. Modelling and simulation of signal transductions in an apoptosis pathway by using timed Petri nets. *Journal of Biosciences*, 32:113–127, 2007.
- H. Matsuno, Y. Tanaka, H. Aoshima, A. Doi, M. Matsui, and S. Miyano. Biopathways representation and simulation on hybrid functional Petri net. In *Proceedings of In Silico Biol. 3*, pages 389–404. IOS Press, 2003.
- T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, April 1989.
- J.S. Oliveira. An algebraic-combinatorial model for the identification and mapping of biochemical pathways. *Bulletin of Mathematical Biology*, 63:1163–1196, 2001.
- D. Purves, G.J. Augustine, and D. Fitzpatrick. *Neuroscience*. Sinauer, Hampshire, UNITED KINGDOM, 1997. ISBN 0878937250.
- V.N. Reddy, M.L. Mavrovouniotis, and M.N. Liebman. Petri net representations in metabolic pathways. In *ISMB-93 Proceedings.*, 1993.
- J.J. Tyson, K.C. Chen, and B. Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15:221–231, 2003.
- C.R. Vazquez, J.H. van Schuppen, and Manuel Silva. A modular-coordinated control for continuous petri nets. In *Proceedings of the 18th IFAC World Congress*, pages 6029–6035, Milano, 2011. IFAC.